

Gene expression analysis

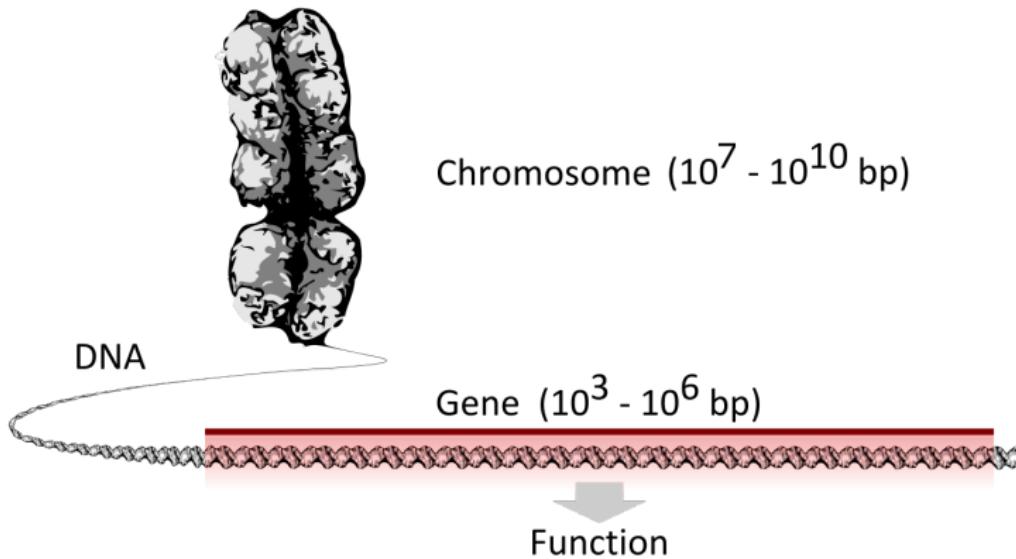
BIOTECH-7005-BIOINF-3000 (2022): Bioinformatics and Systems
Modelling

Zhipeng Qu

School of Biological Sciences,
The University of Adelaide

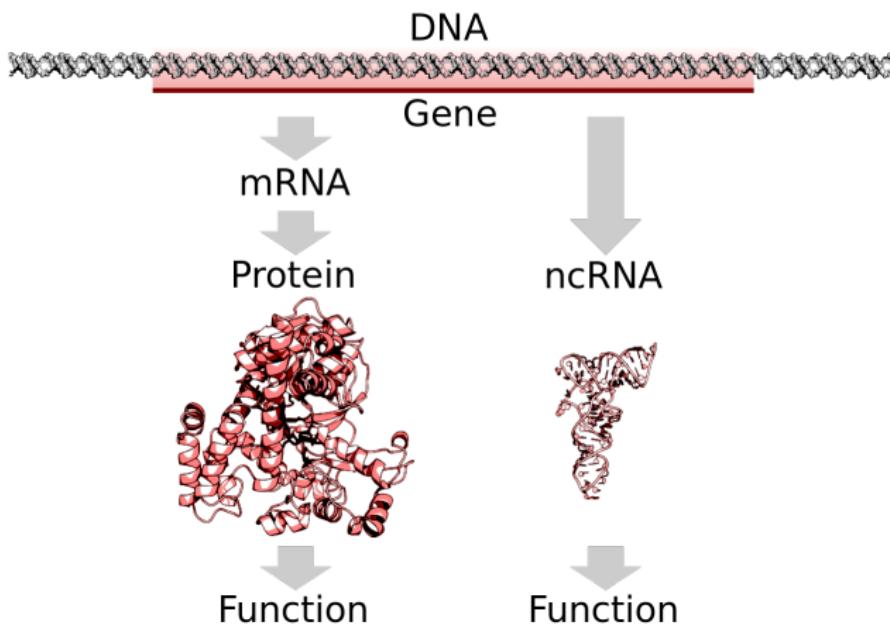
October 10th, 2022

What is a gene



"A gene is a region of DNA that encodes function. A chromosome consists of a long strand of DNA containing many genes. A human chromosome can have up to 500 million base pairs of DNA with thousands of genes."

Gene expression

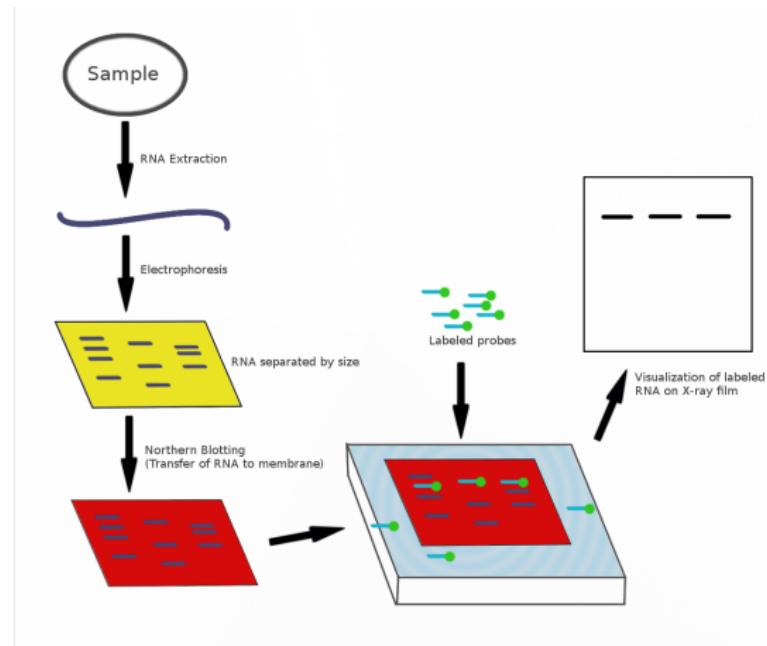


Gene expression is measured at RNA level.

Some approaches used for quantifying gene expression

- Northern blotting
- qRT-PCR
- Microarray
- RNA-Seq
- others...

Northern blotting



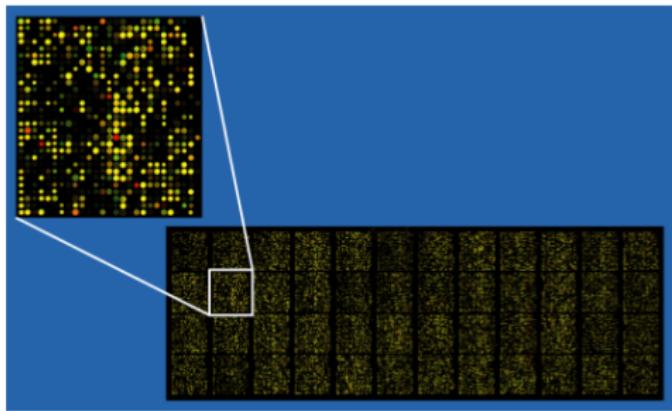
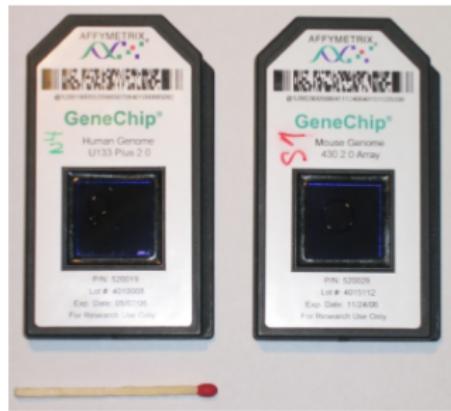
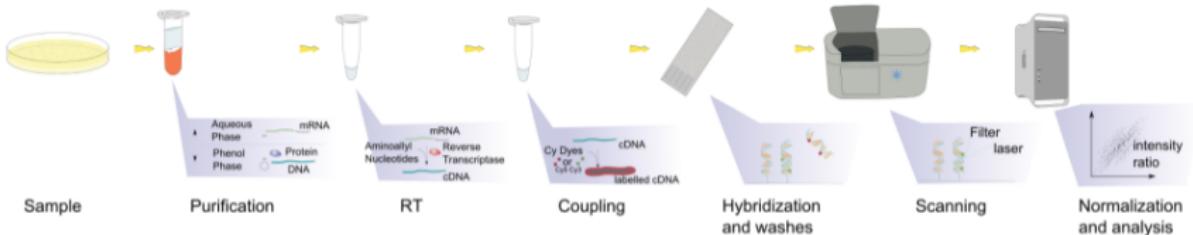
Northern blotting takes its name from its similarity to the first blotting technique, the Southern blot (DNA detection).

https://en.wikipedia.org/wiki/Northern_blot

qRT-PCR

Technique	Abbreviation
Polymerase chain reaction	PCR
Reverse transcription polymerase chain reaction	RT-PCR
Real-time polymerase chain reaction	qPCR
RT-PCR / qPCR combined technique	qRT-PCR

Microarray



“RNA-Seq, also called RNA sequencing, is a particular technology-based sequencing technique which uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment, analyzing the continuously changing cellular transcriptome.”

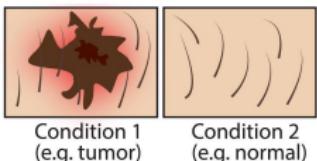
<https://en.wikipedia.org/wiki/RNA-Seq>

Wang Z, Gerstein M, Snyder M (January 2009). "RNA-Seq: a revolutionary tool for transcriptomics". Nature Reviews. Genetics. 10 (1): 57–63.

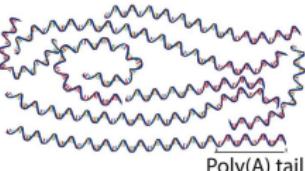
Overview of RNA-Seq

Part 1, Library preparation

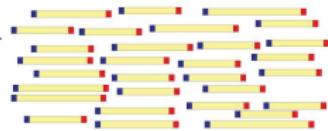
Samples of interest



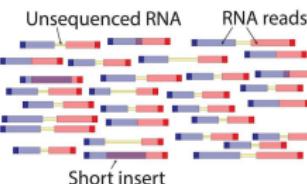
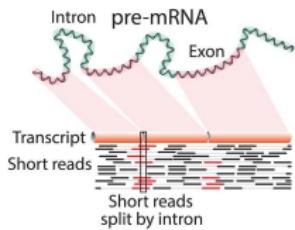
Isolate RNAs



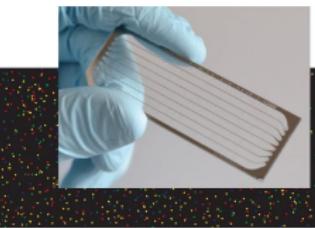
Generate cDNA, fragment, size select, add linkers



Map to genome, transcriptome, and predicted exon junctions



Sequence ends

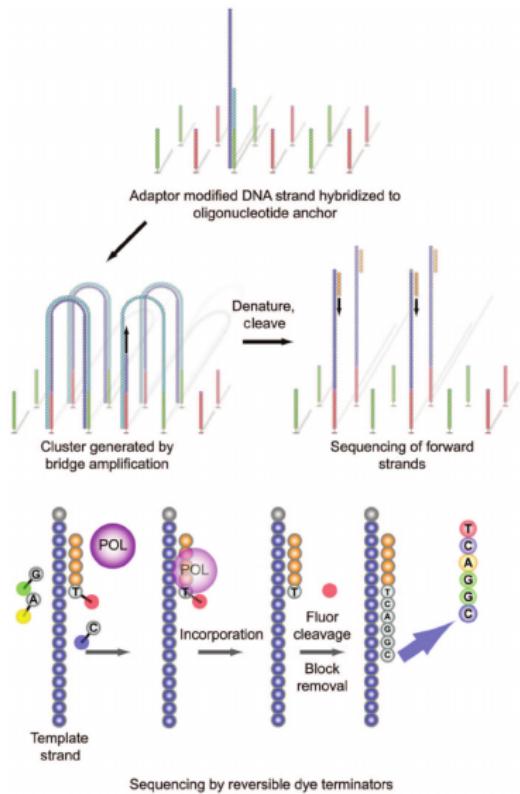


Downstream analysis

Part 3, Bioinformatics analysis and Downstream analysis

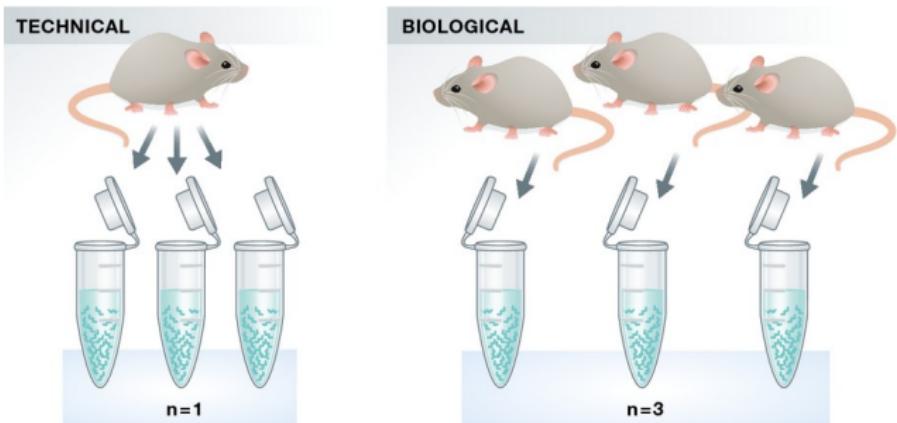
Part 2, Next generation sequencing

Illumina sequencing



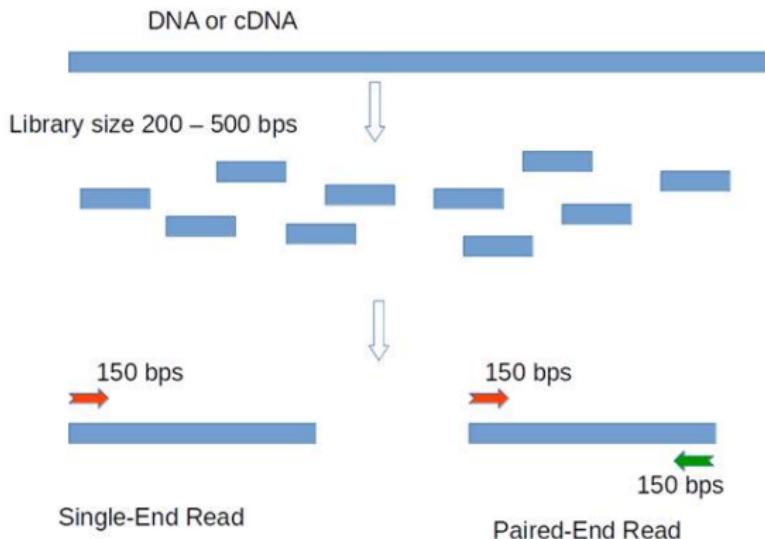
Karl V Voelkerding, Shale A Dames, Jacob D Durtschi, Next-Generation Sequencing: From Basic Research to Diagnostics, Clinical Chemistry, Volume 55, Issue 4, 1 April 2009.

Technical replicates VS biological replicates



- **Technical replicates** use the same biological sample to repeat the technical or experimental steps in order to accurately measure technical variation and remove it during analysis.
 - **Biological replicates** use different biological samples of the same condition to measure the biological variation between samples.

Paired end VS single end



- Single End (SE): only one end of each cDNA fragment is sequenced.
- Paired End (PE): both ends of each cDNA fragment are sequenced. Sequencing reads are labeled as pairs.

Sequencing depth in RNA-Seq

Sequencing depth in RNA-Seq is the number of reads in each sample. It varies depending on the goal of the study.

- Normal gene expression profiling experiments: 5–25 million reads per sample
- Experiments looking for a more global view of gene expression, or alternative splicing: 30–60 million reads per sample
- In-depth view of transcriptome, or assemble new transcripts: 100–200 million reads per sample
- Targeted RNA-Seq: 3 million reads per sample
- Small RNA-Seq: 1–5 million reads per sample

Read length in RNA-Seq

Read length will depend on the application and final size of the library.

- Gene expression profiling: SE50–SE75
- Transcriptome annotation or assembly: longer, paired-end reads (such as 2 x 75 bp) to enable more complete coverage of the transcripts and identification of novel variants or splice sites
- small RNA-Seq: SE50

https://sapac.support.illumina.com/bulletins/2017/04/considerations-for-rna-seq-read-length-and-coverage-.html

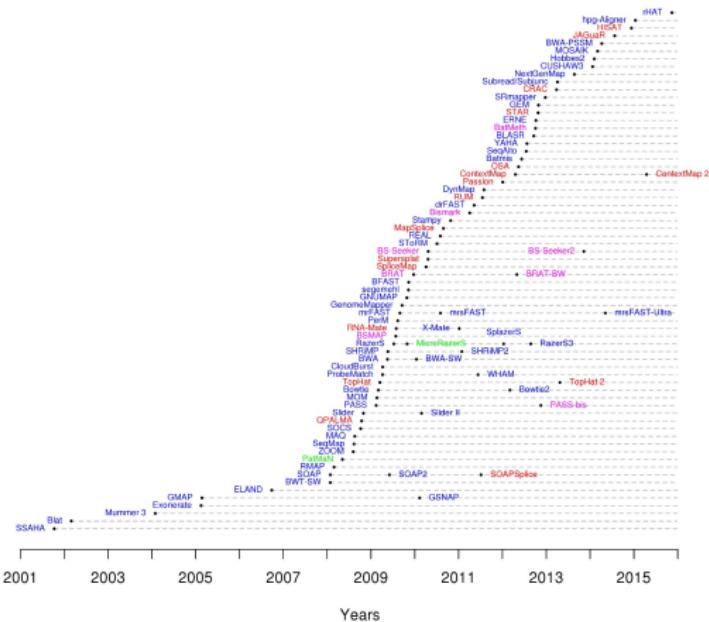
Short read alignment

Also called read mapping, align (map) short reads from NGS to reference genome (if available, DNA-Seq/RNA-Seq) or transcriptome (RNA-Seq).

Challenges in RNA-Seq alignment:

- millions of short reads (DNA-Seq/RNA-Seq)
- RNA splicing

Short read alignment tools (short aligners)



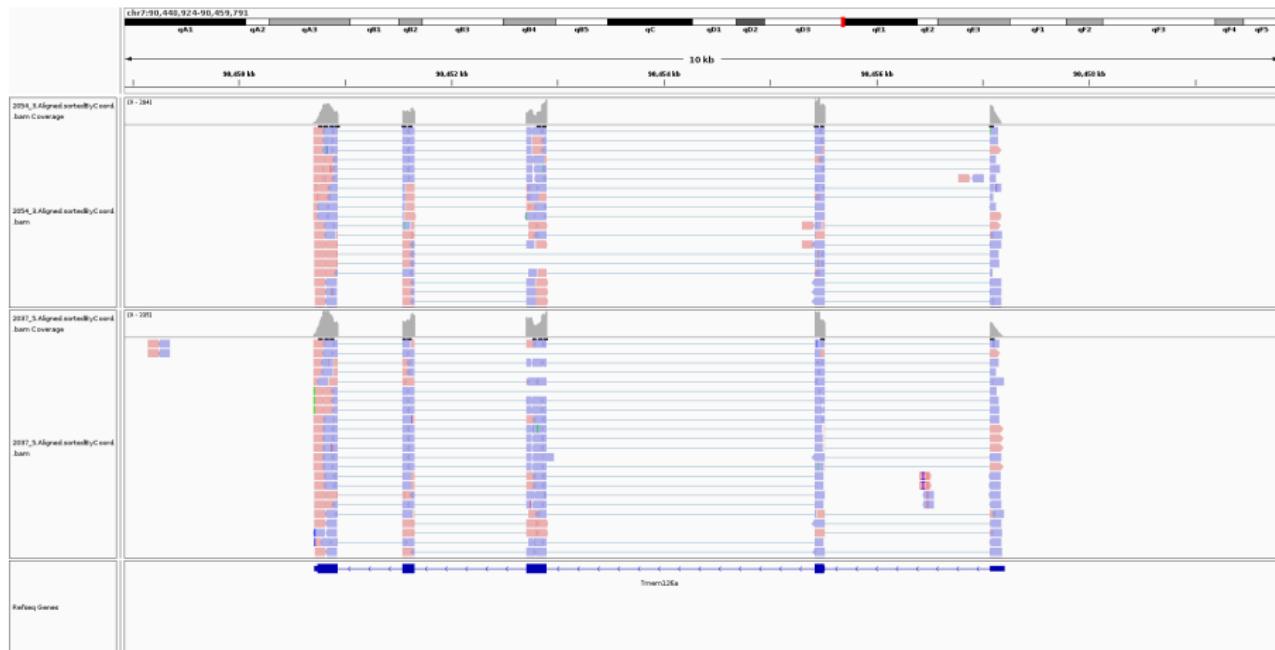
Which short aligner should I use?

- Sequencing input type: DNA vs RNA
- Reference sequences: Genome vs Transcriptome
- Available computing resources

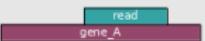
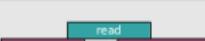
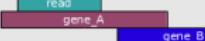
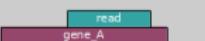
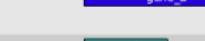
STAR (Spliced Transcripts Alignment to a Reference)

- Outperforms other aligners by more than a factor of 50 in mapping speed
- Memory intensive. At least 10x Genome size (for example, ~30 Gb for human genome)
- Written in C++, only works on Linux or Mac OS
- Unbiased de novo detection of canonical junctions
- Discover non-canonical splices and fusion transcripts

Reads mapped to gene



Three read count modes

	union	intersection _strict	intersection _nonempty
 A single read aligned to gene_A.	gene_A	gene_A	gene_A
 A single read partially aligned to gene_A.	gene_A	no_feature	gene_A
 A single read overlapping gene_A.	gene_A	no_feature	gene_A
 Two reads aligned to gene_A.	gene_A	gene_A	gene_A
 A single read aligned to both gene_A and gene_B.	gene_A	gene_A	gene_A
 A single read aligned to both gene_A and gene_B.	ambiguous (both genes with --nonunique all)	gene_A	gene_A
 A single read aligned to both gene_A and gene_B.	ambiguous (both genes with --nonunique all)		
 A single read aligned to both gene_A and gene_B. The alignment is marked as ambiguous.	alignment_not_unique (both genes with --nonunique all)		

Multiple mapping

Short reads can be mapped to multiple features (genes/transcripts)

- Identical/similar sequences in different genes (e.g. gene family, repetitive elements)
- Different transcription isoforms from same gene

Species	Aligner	Read length	multiple mapping rate (%)
Human	STAR	PE100	4.88
Mouse	STAR	PE100	15.72
Rat	STAR	PE75	12.07
Arabidopsis	STAR	PE150	1.41
Rice	Tophat2	PE150	43.7
Soybean	Tophat2	PE150	26.4

Strategies for handling multiple mapping

- Use uniquely mapping reads only
- Simple “rescue” method. Uniformly divide each multi-mapping read to all of the positions it maps to. In other words, a read mapping to 10 positions will count as 10% of a read at each position.
- “Rescue” method using Expectation-Maximization model
 - ① E-step (Expectation) Give transcript abundances, estimate the probability of each read mapping to each transcript
 - ② M-step (Maximization) Update the abundances by redistributing the reads
 - ③ Go to step 1 (E-step) until convergence

RNA-Seq is a relative abundance measurement of RNA expression level

- Short reads are RNA fragments randomly picked and sequenced from library
- Additional information, such as levels of “spike-in” transcripts, are required for absolute measurements
- Normalization of read count is needed to compare gene/transcript abundance
 - ① RPKM/FPKM (Reads/Fragments Per Kilobase Million)
 - ② TPM (Transcripts Per Million)

RPKM vs TPM

- RPKM/FPKM
 - Corrects for total library coverage
 - Corrects for gene length
 - Comparable between different genes within the same dataset
 - FPKM can be used for paired end libraries
- TPM
 - Normalises to transcript copies instead of reads
 - corrects for cases where the average transcripts length differs between samples

Identification of DE genes

The diagram illustrates a heatmap of gene expression data. The rows represent genes, and the columns represent biological replicates. The genes are grouped by a bracket labeled "Genes". The replicates are grouped into two main categories: "Wild type" (blue bracket) and "Mutant" (red bracket). Within each category, replicates are further grouped by a bracket labeled "Biological replicates". The replicates are labeled as follows: Wild type (2074-1, 2136-2, 2220-4, 2167-8), and Mutant (2178-2, 2178-1). The data values are represented by numbers in the cells of the heatmap.

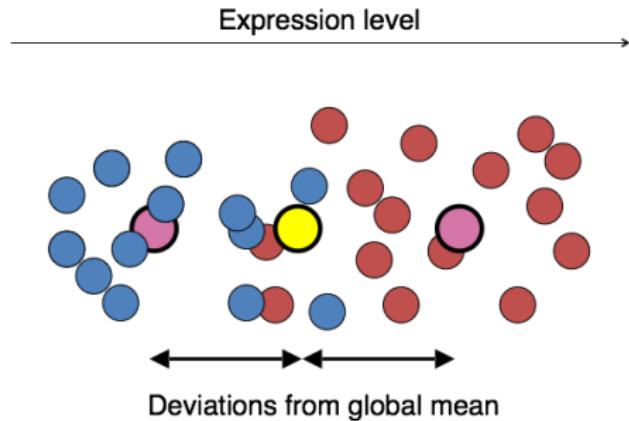
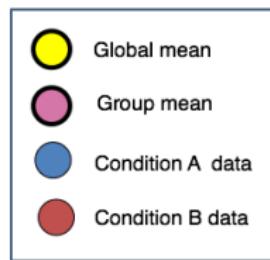
gene_name	Wild type			Mutant		
	2074-1	2136-2	2220-4	2167-8	2178-2	2178-1
4933401J01Rik	0	0	0	0	0	0
Rp1	0	1	0	1	1	0
Gm6101	0	0	0	0	0	0
Gm37483	0	0	0	0	0	0
Sox17	241	145	229	359	314	276
Gm37587	0	0	0	1	0	1
Gm7357	1	3	3	3	8	4
Gm22307	0	0	1	0	2	0
Gm38076	0	0	0	0	0	0
Gm25493	2	1	0	0	1	2
Gm2053	0	0	1	0	0	0
Gm6123	8	26	30	26	32	23
Mrpl15	434	419	417	472	439	408
Gm37144	2	1	0	0	1	0
Lypla1	571	493	445	594	818	648
Gm37988	0	0	0	0	0	0
Tcea1	513	544	504	554	515	490
...						

Identification of DE genes

Genes	Wild type			Mutant			Up? Down?
	2074-1	2136-2	2220-4	2167-8	2178-2	2178-1	
gene_name							
4933401J01Rik	0	0	0	0	0	0	
Rp1	0	1	0	1	1	0	
Gm6101	0	0	0	0	0	0	
Gm37483	0	0	0	0	0	0	
Sox17	241	145	229	359	314	276	
Gm37587	0	0	0	1	0	1	
Gm7357	1	3	3	3	8	4	
Gm22307	0	0	1	0	2	0	
Gm38076	0	0	0	0	0	0	
Gm25493	2	1	0	0	1	2	
Gm2053	0	0	1	0	0	0	
Gm6123	8	26	30	26	32	23	
Mrpl15	434	419	417	472	439	408	
Gm37144	2	1	0	0	1	0	
Lypla1	571	493	445	594	818	648	
Gm37988	0	0	0	0	0	0	
Tceal1	513	544	504	554	515	490	
...							

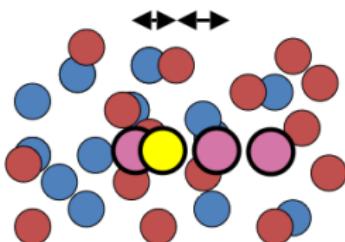
https://github.com/hbctraining/rnaseq-cb321/blob/master/lessons/analysis_methods.md

Identification of DE genes



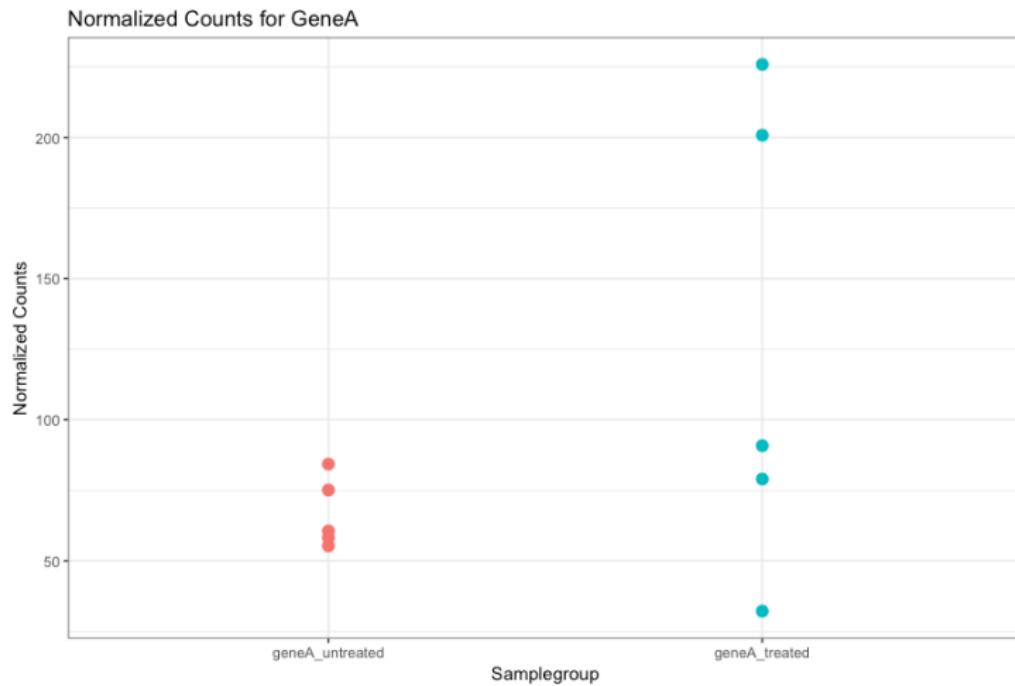
Significant difference

Deviations from global mean

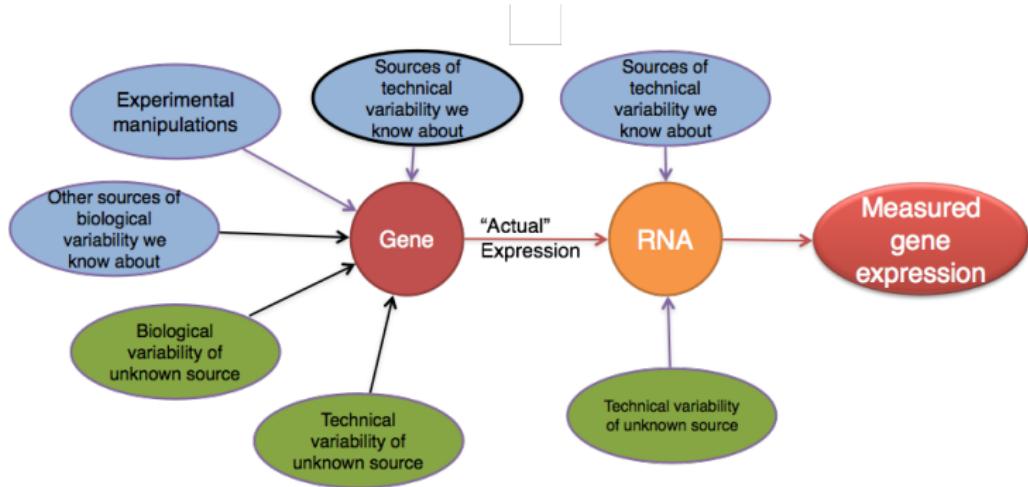


No significant difference

Identification of DE genes



Variability on gene expression

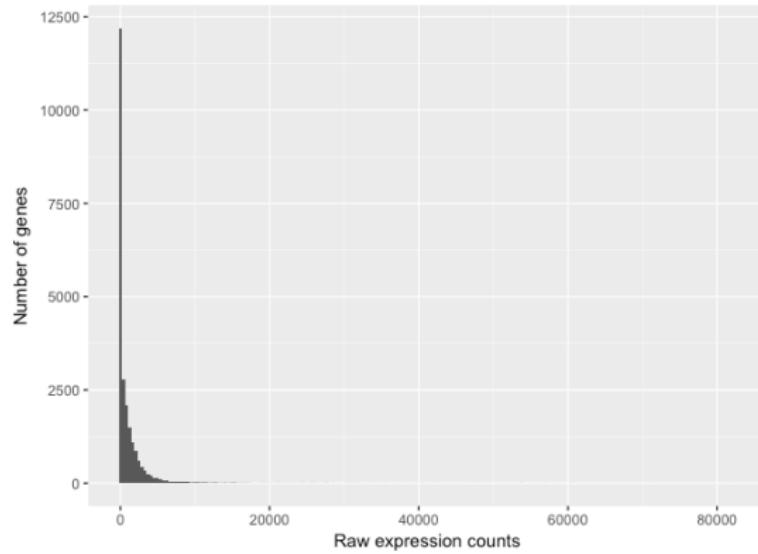


Courtesy of Paul Pavlidis, UBC

The goal of differential expression analysis is to identify and correct for sources of variation such that we can separate the “interesting” from the “uninteresting”.

Properties of gene expression data from RNA-Seq

- Many non-/low- expressed genes (0s in the raw count table)
 - Large dynamic range for the expression value
 - Clearly not following normal distribution



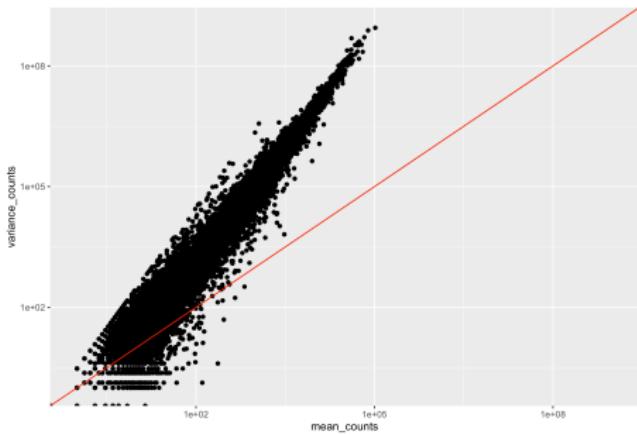
Models/tools used in DE analysis

- Poisson distribution
- Negative Binomial (NB) distribution

Can we use Poisson distribution to model gene expression data?

- Poisson distribution: For use, when the number of cases is very large (i.e. people who buy lottery tickets), but the probability of an event is very small (probability of winning).
- With RNA-Seq data, a very large number of RNAs are represented and the probability of pulling out a particular transcript is very small.
- Poisson distribution has **the same mean and variance**.

Mean VS variance on gene expression data



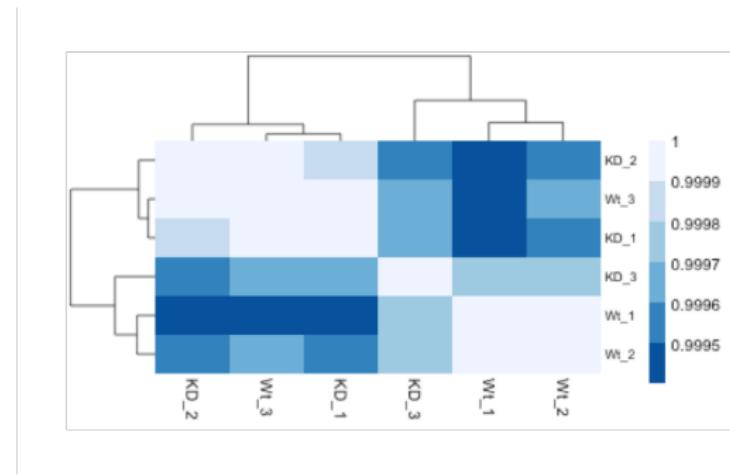
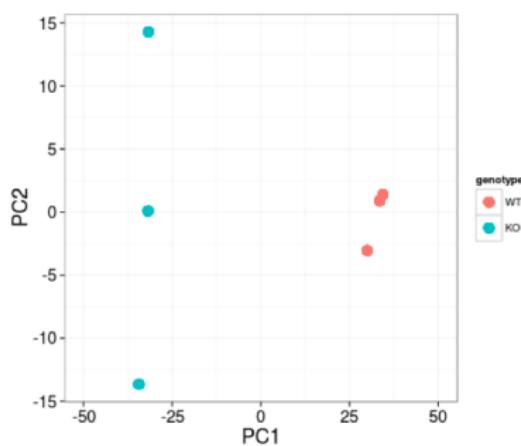
- We can use the scatter plot to show the relationship between mean values and variance values in a sample
 - Mean is not equal to the variance
 - Poisson distribution does not fit the gene expression data very well when we have small number of replicates (what else?)

Negative Binomial (NB) distribution

- Example: Gives you the probability of getting a number of heads upon tossing a coin a number of times. Based on discrete events and used in situations when you have a certain number of cases. Thus, it gives the probability of getting r events out of n . trials.
- NB is a good approximation for data where the mean < variance.
- DE analysis tools using NB model:
 - edgeR (Robinson et al, 2010)
 - DESeq2 (Love et al, 2014)

Assess overall similarity between samples (sample-level QC)

- How similar are biological replicates from the same sample?
 - How different are samples from different groups (e.g. Mutant vs WT)
 - Is the experimental condition the major source of variation?



What you will get from a typical DE analysis

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.01098240411
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.01098240411
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Gramd2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000005220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

What you will get from a typical DE analysis

Gene name

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.01098240411
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.01098240411
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Grand2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000052220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

What you will get from a typical DE analysis

log(Mutant/WT)

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.01098240411
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.01098240411
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Gramd2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000005220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

What you will get from a typical DE analysis

Average expression

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.01098240411
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.01098240411
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Gramd2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000005220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

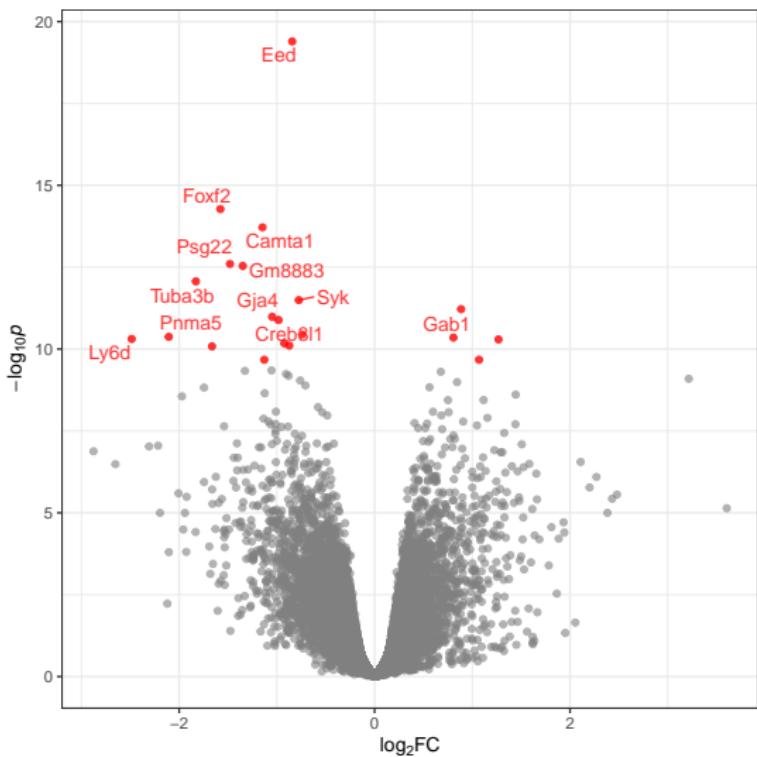
What you will get from a typical DE analysis

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.01098240411
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.01098240411
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Gramd2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000005220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

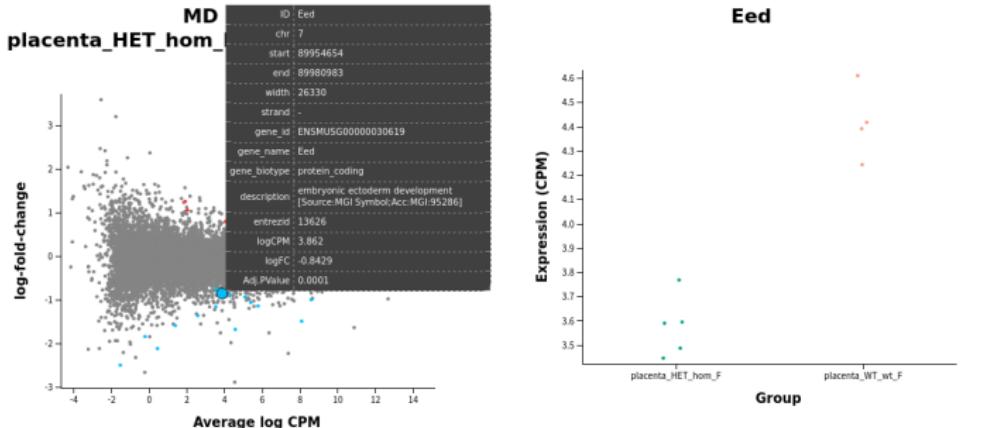
What you will get from a typical DE analysis

gene_id	gene_name	gene_biotype	logFC	AveExpr	P.Value	Adjusted P value (FDR)
						adj.P.Val
ENSMUSG00000030619	Eed	protein_coding	-0.8429485747	3.861989267	3.78E-09	5.79E-05
ENSMUSG00000038402	Foxf2	protein_coding	-1.578162112	1.378502208	6.32E-07	0.004837630789
ENSMUSG00000014592	Camta1	protein_coding	-1.147416231	3.528201278	1.10E-06	0.005640100289
ENSMUSG00000044903	Psg22	protein_coding	-1.479941884	8.085570643	3.37E-06	0.010982404111
ENSMUSG00000099707	Gm8883	lincRNA	-1.349036399	2.575934167	3.59E-06	0.010982404111
ENSMUSG00000067338	Tuba3b	protein_coding	-1.829221491	-0.224905613	5.74E-06	0.01466066189
ENSMUSG00000021457	Syk	protein_coding	-0.7752655801	4.002586886	1.02E-05	0.02228335849
ENSMUSG00000031714	Gab1	protein_coding	0.8860410024	7.125652031	1.34E-05	0.02569841359
ENSMUSG00000050234	Gja4	protein_coding	-1.047571965	5.385033109	1.69E-05	0.02863564488
ENSMUSG00000027230	Creb3l1	protein_coding	-0.9828163804	8.595367353	1.87E-05	0.02863564488
ENSMUSG00000052852	Reep1	protein_coding	-0.7428457788	3.548621004	2.91E-05	0.03453646801
ENSMUSG00000050424	Pnma5	protein_coding	-2.106134584	0.435498476	3.12E-05	0.03453646801
ENSMUSG00000074259	Gramd2	protein_coding	0.8069535376	4.042051003	3.20E-05	0.03453646801
ENSMUSG00000034634	Ly6d	protein_coding	-2.485929445	-1.537158461	3.33E-05	0.03453646801
ENSMUSG00000043089	Mmp1a	polymorphic_pseudogene	1.2682024	1.908758712	3.38E-05	0.03453646801
ENSMUSG00000005220	Corin	protein_coding	-0.9235405724	5.077351708	3.81E-05	0.03553531432
ENSMUSG00000050271	Prag1	protein_coding	-0.8729805051	4.294968795	4.08E-05	0.03553531432
ENSMUSG00000023046	Igfbp6	protein_coding	-1.663887879	4.563679516	4.18E-05	0.03553531432
ENSMUSG00000085860	2410003L11Rik	processed_transcript	1.069083051	2.020513905	6.29E-05	0.04817316657
ENSMUSG00000037185	Krt80	protein_coding	-1.129170053	5.773404496	6.29E-05	0.04817316657

What you will get from a typical DE analysis - Volcano plot



What you will get from a typical DE analysis - Glimma plot



ID	chr	start	end	width	strand	gene_id	gene_name	gene_biotype	description
Eed	7	89954654	89980983	26330	-	ENSMUSG0000030619	Eed	protein_coding	embryonic ectoderm development [Source:MGI Symbol;Acc:MGI:95286]
Foxf2	13	31625816	31631403	5588	+	ENSMUSG0000038402	Foxf2	protein_coding	forkhead box F2 [Source:MGI Symbol;Acc:MGI:1347479]
Camta1	4	150917322	151861876	944555	-	ENSMUSG0000014592	Camta1	protein_coding	calmodulin binding transcription activator 1 [Source:MGI Symbol;Acc:MGI:2140230]

Thank you!